

Nós na rede. Conexão e participação em dois grupos do Facebook a partir dos digital methods

Jorge Martins Rosa

dedalus.jmmr@gmail.com

Universidade Nova de Lisboa

Janna Joceli C. de Omena

jannajoceli@gmail.com

Universidade Nova de Lisboa

Resumo

Na senda das BBS e da Usenet, os grupos do Facebook são atualmente, por um efeito de arrasto da popularidade desta rede social *online*, locais privilegiados de discussão orientada em torno de temas e interesses comuns. Apesar da sua omnipresença, e ao contrário dos perfis individuais e das páginas institucionais, a análise dos seus fluxos de interação tem ocupado um lugar secundário na pesquisa académica. E nesta os métodos mais clássicos, mesmo que adaptados ao meio (da etnografia para a netnografia, por exemplo), têm dominado. O que se propõe aqui é, recorrendo a ferramentas de recolha e visualização de dados, de que são exemplo o Netvizz e o Gephi, demonstrar o modo como as metodologias computacionais, inspiradas na Teoria das Redes e na Análise de Redes Sociais, potenciam práticas de pesquisa que, ao apresentarem uma homologia com os próprios meios que são o seu objeto, permitem – tomando os utilizadores e suas acções (*posts* e comentários) como nós de um grafo – recolher informação tanto ou mais repleta de significado acerca do modo como aí se interage. Tomaremos como estudo de caso dois grupos, respetivamente dedicados à partilha e discussão de música e de cinema, focando-nos – com o auxílio dessas ferramentas – no modo como aí participam e interagem. Recorrendo a diversas ferramentas dos Métodos Digitais, procura-se, de forma exploratória, apurar quais as variáveis relevantes para a compreensão destes casos específicos, nomeadamente no que respeita à existência de utilizadores mais influentes, formulando hipóteses para futuras análises mais aprofundadas.

Palavras-chave: métodos digitais; Facebook; exploratory data analysis; grupos; envolvimento; teoria das redes.

1. Introdução

Se no final da década de 90, com aquilo a que podemos chamar a primeira geração dos «Internet Studies», um dos conceitos-chave era o de «comunidade»¹, hoje em dia este parece

1. Cf. títulos de colectâneas como *Cybersociety: Computer-Mediated Communication and Community* e *Cybersociety 2.0: Revisiting Computer-Mediated Community and Technology* (ambos organizados por Steven G. Jones), ou *Communities in Cyberspace* (organizado por Peter Kollock e Mark Smith), e ainda o clássico de divulgação *The Virtual Community: Homesteading on the Electronic Frontier*, de Howard Rheingold.

estar algo adormecido² apesar da ascensão – para não dizer triunfo – das plataformas ou *sites* de redes sociais [SRS], com o Facebook inegavelmente na dianteira. E, com essa ascensão, também a de «rede» enquanto conceito operacional, cremos que em boa parte – e também este artigo procura integrar-se em tal tendência – devido a uma atenção crescente ao contributo da Teoria das Redes e das metodologias que lhe são afins.

Contudo, ainda que possamos aplaudir a maior objetividade e rigor que decorrem dessa viragem quantitativa e computacional, bem como a preferência por um conceito também mais neutro como é o de «rede», a ideia de que estas plataformas podem servir como facilitadores de uma interação regular, e portanto de coesão, entre os seus utilizadores³ parece também estar a ser preterida em favor ora de abordagens em que o indivíduo é o foco (modalidades de utilização, questões de privacidade, etc.), ora de outras em que este é pouco mais do que um «seguidor» ou «cliente» de páginas institucionais e comerciais. Estamos, como entretanto deve ter-se tornado claro, a referir-nos às três principais funcionalidades do Facebook enquanto espaço de expressão⁴: os perfis individuais, as páginas (maioritariamente institucionais, distribuídas por diversas categorias⁵, embora exista a reveladora opção genérica de classificação como «community page»), e ainda aos grupos. É nesta última funcionalidade, precisamente aquela com mais semelhanças com as «comunidades» que tanto monopolizavam os artigos do final da década de 90, que identificamos um declínio de atenção por parte da academia, que aqui queremos contrariar.

2. Uma plataforma, uma rede, dois grupos

Somos os primeiros a reconhecer que é nos grupos que por um lado se encontram as maiores barreiras a qualquer tarefa de investigação, pois enquanto as páginas são por predefinição públicas estes podem, além de abertos, também ser privados ou secretos⁶, e portanto de mais difícil acesso a não ser que o investigador neles se integre, como observador (ora mais ora menos) par-

2. Não é esta a ocasião mais adequada para uma confirmação sustentada dessa aparente tendência, pelo que nos limitamos a constatar que uma busca por «community “social network”» no Google Scholar apenas devolve, nas primeiras cinco páginas, cinco resultados em que simultaneamente «community» surge no título e é inequívoco que «social network» se refere a um SRS: um artigo de 1999 sobre o IRC («The Virtual Speech Community: Social Network and Language Variation on IRC», de John Paolillo), o ultra-citado «Friends, Friendsters, and Myspace Top 8: Writing Community into Being on Social Network Sites», de dannah boyd (de 2006), «Word of Mouth Communication within Online Communities: Conceptualizing the Online Social Network», de Jo Brown, Amanda J. Broderick e Nick Lee (de 2007), «An Evaluation of Identity-Sharing Behavior in Social Network Communities», de Frederic Stutzman (de 2006), e o livro *A Networked Self: Identity, Community, and Culture on Social Network Sites*, organizado por Zizi Papacharissi.

3. Note-se como tentámos propositadamente evitar o termo «comunidade».

4. Faltam pelo menos mais duas funcionalidades, transversais a essas três principais, que deixamos aqui de lado. Uma é a das aplicações, de certa forma porque diluem as fronteiras do Facebook enquanto site (que, por exemplo, através da comunicação entre as respetivas APIs permitem aí replicar publicações feitas no Twitter ou no Instagram). A outra é a da criação e divulgação de eventos.

5. Figura pública, organização, causa, músico/banda, organização política, etc.

6. Ambos com entrada sujeita a aprovação dum administrador, mas visíveis na busca se privados (e portanto sendo possível um pedido de adesão), e invisíveis se secretos (e portanto dependentes de um convite de outro membro ou inclusão direta por parte dum administrador).

ticipante. E, por outro lado, pode também presumir-se uma maior diversidade – para não dizer imprevisibilidade – de dinâmicas de grupo para grupo, dada a inumerável quantidade de motivos subjacentes à sua criação e continuação⁷, que no limite poderão impedir qualquer tentativa de generalização. Conscientes disso, e apesar de terminarmos com uma hipótese de caráter global, declinamos portanto qualquer pretensão de generalizar tudo o que abaixo se segue, pouco mais havendo do que uma forte curiosidade em torno da sua dinâmica diária a determinar a escolha dos dois grupos que são objeto deste artigo.

É por isso necessário caracterizá-los, nem que seja de forma breve. Num artigo anterior (Rosa, 2015), foi tentada uma primeira abordagem – já com alguns elementos quantitativos, mas maioritariamente qualitativa – a um deles, com uma curtíssima menção ao outro. Começamos pelo que têm em comum. Ambos abrigam na sua gênese paixões, ora pela música – no grupo intitulado «Igreja Universal dos Fazedores de Bonitas Listas Musicais dos Últimos Dias»⁸ – ora pelo cinema – na «Tribo Cinéfila das Tapas e Vinho Tinto»⁹, e essas paixões são mobilizadas (ou seja, é incentivada a participação) da mesma forma: todos os dias um dos administradores propõe um tema e os membros (administradores incluídos) podem publicar uma e só uma ligação (quase invariavelmente para um vídeo no YouTube, como se verá adiante) apontando ora para uma música ora para um excerto de filme que cumpra esse mesmo tema diário¹⁰. Esta semelhança não é fruto do acaso: o primeiro dos grupos, com cerca de 14 mil membros em novembro de 2015, foi criado a 28 de novembro de 2009; o segundo, que surgiu a 7 de junho de 2010 e tinha no mesmo mês pouco mais de 2700 membros, foi um seu *spin-off* e partilha com este muitos membros e também administradores.

Mesmo que nos limitássemos a este retrato, salta à vista uma primeira diferença, que aliás obrigou a alguns ajustes metodológicos na nossa pesquisa (cf. abaixo): os cerca de sete meses de atraso no início de atividades na «tribo» não são suficientes para justificar tão grande discrepância no número de membros, que nem deve ser contabilizada em valores absolutos mas sim relativos – aproximadamente 5,6 vezes mais na «igreja». Contudo, deixando já pistas para a análise que se segue, será também prematuro afirmar que se trata apenas de uma questão de proporcionalidade.

2.1. Uma abordagem exploratória

Esclareçamos então, como advertência prévia, qual é a postura metodológica desta análise ao funcionamento de ambos os grupos, bem como a dos dois investigadores e autores envolvidos. Ambos somos membros dos dois grupos, num caso por razões exteriores (e temporalmente anteriores) à oportunidade da pesquisa, no outro resultando como necessidade desta, e o grau

7. Apenas pela experiência quotidiana enquanto utilizadores do Facebook, e de forma avulsa, podemos mencionar grupos que congregam atuais e antigos colegas de cursos, adeptos de clubes ou praticantes de algum desporto, classes profissionais como a dos jornalistas, de debate político, de ficção (literária ou outra, restrita ou não a algum género), de fãs de celebridades, etc. E mesmo entre grupos cujo «centro de atração» é similar, as atividades de cada um podem variar a ponto de desencorajar qualquer comparabilidade.

8. Cf. www.facebook.com/groups/228544121320/. Doravante, «igreja» (com aspas) ou IUFBLMUD.

9. Cf. www.facebook.com/groups/tcdtevt/. Doravante, «tribo» (com aspas) ou TCTVT.

10. Em Rosa, 2015 é feita uma descrição mais detalhada do funcionamento da IUFBLMUD e respetivas regras, que no essencial se repetem para a TCTVT.

de participação está em consonância: um participante frequente e ativo, entretanto promovido a administrador de um dos grupos; uma observadora com participações muito esporádicas. Trata-se por agora de uma pesquisa exploratória guiada por dados, ou «exploratory data analysis» (cf. Tukey, 1977), com recurso aos «Digital Methods»¹¹, que procura trazer para primeiro plano algumas intuições que antecediam a recolha de dados, e – de um modo quase inevitável nesta fase – ainda repleta de especulações na leitura dos mesmos. Em vez de uma hipótese prévia a guiar a pesquisa, mas esperando que ela possa dar origem a um projeto de investigação mais extenso e mais profundo, deixou-se antes que os valores, as métricas e inclusive alguns gráficos tivessem um estatuto heurístico, gerando de forma orgânica futuras hipóteses muito mais do que respondendo mecanicamente a outras que estariam a ser-lhes impostas.

2.2. Metodologia

Os dados que alimentam a análise que se segue foram obtidos através do Netvizz, aplicação *online* que solicita à API do Facebook, e de acordo com o que esta autoriza, um conjunto de dados relativos a publicações, utilizadores, *likes* e comentários, e os estrutura em dois tipos de documentos, para tratamento posterior: CSV (compatível com Excel) e GDF (para análise e representação visual da rede, por exemplo através do programa Gephi). Numa primeira recolha preliminar, a 6 de julho de 2015, foram selecionados os 999 *posts* mais recentes de cada grupo, o que desde logo permitiu confirmar uma intuição prévia respeitante à disparidade no grau de atividade em cada um dos grupos. O mesmo número de publicações correspondia, no caso da «tribo», a um período ligeiramente superior a um mês e a 311 utilizadores ativos (isto é, publicando, fazendo *like* ou comentando), enquanto na «igreja» um número equivalente de publicações foi atingido em apenas sete dias, contabilizando-se em mais de mil o número de utilizadores ativos. Perante esta discrepância, e de forma a garantir uma mais adequada comparação entre as dinâmicas de cada um dos grupos, o processo de extração de dados foi repetido, mas desta vez mantendo fixo o maior período de tempo, em dias inteiros, encontrado na primeira tentativa, isto é, as datas inicial e final da «tribo»: de 30 de maio a 6 de julho de 2015.

A análise aos dados recolhidos, que a seguir se discrimina, pode ser categorizada da seguinte forma: i) visão geral de cada grupo e caracterização dos seus membros; ii) análise das publicações de acordo com o seu tipo e envolvimento («engagement») gerado; iii) análise da rede de interação entre membros e publicações.

3. Visão geral da atividade dos grupos

3.1. A Tribo Cinéfila das Tapas e Vinho Tinto

Como anunciado acima, a aplicação Netvizz¹² foi utilizada para recolher os dados de atividade dos grupos durante o período indicado. Para este grupo de menores dimensões, chegou-se

11. Os Métodos Digitais (Rogers, 2013) pressupõem que é possível tomar a *web*, e em particular as redes sociais, simultaneamente pelo seu valor intrínseco como objetos de pesquisa e pelo seu valor de fonte para compreender a realidade social de onde emanam.

12. Cf. <https://apps.facebook.com/netvizz/>

a um total de 982 publicações¹³, sendo 311 o número de utilizadores ativos – isto é, que publicaram, gostaram ou comentaram *posts*, somando estas três ações possíveis um total de 14.028.

É inevitável que nos interroguemos por um lado se ao longo dos dias esta atividade é uniforme, e por outro, talvez ainda mais relevante – quanto mais não seja por se verificar que, de entre os cerca de 2700 membros¹⁴, apenas um pouco mais de 10% dos utilizadores é ativo –, se estes que restam se assemelham ou não no seu grau de atividade. A tabela que se segue, distribuindo o grau de atividade por quartis (e acrescentando o decil superior), mostra que essa uniformidade tende a sofrer desvios significativos.

No que respeita às publicações, são 982 distribuídas ao longo de 38 dias, o que dá uma média (arredondada às unidades) de 26 por dia. Este número oscila entre as 8 publicações no dia 21 de junho e as 43 a 1 de junho, o que aponta para bastante variabilidade, mas tomados os quartis podemos verificar que esta não se traduz em irregularidade: 21 *posts* representam o quartil inferior, 26 a mediana (praticamente coincidindo com a média), e 36 o quartil superior. Se consideradas as outras atividades que revelam envolvimento (comentários e *likes*, visíveis também no Quadro 1), o panorama em parte repete-se, embora – como veremos com mais pormenor logo abaixo – não segundo uma proporcionalidade simples. Isto é, há uma variação, ao longo dos quartis, no número de *likes* que tende a acompanhar o de publicações (um rácio de 9,8 *likes* por *post* no primeiro quartil, de 11,9 na mediana e no quartil superior, e 14,4 no decil superior). Este rácio parece portanto aumentar à medida que a quantidade de publicações diárias também sobe), ocorrendo o mesmo no caso dos comentários, o que nos leva ao que identificaremos como *Hipótese A*: a de que há um efeito de *feedback* positivo; isto é, quanto mais *posts* num dia, maior o envolvimento também para cada uma das publicações nesse mesmo dia¹⁵.

13. Com uma ligeira discrepância relativamente aos 999 da recolha de teste, dado que a unidade de medida passou a ser o número de dias e não de *posts*. Outros fatores, como a variação no número de membros e a eventual existência de algum tipo de atividade retroativa (fazer ou retirar *like* a publicações antigas, por exemplo), podem igualmente contribuir para estas discrepâncias, que contudo podem ser ignoradas se não se traduzirem em alterações significativas aos indicadores.

14. A aplicação usada não permite apurar qual o número de membros que cada grupo tinha no período em análise (período esse em que, de resto, pode ter havido alteração dessa quantidade), mas a observação quotidiana, quer durante quer depois das datas em causa, permite com bastante segurança avançar com esse valor aproximado.

15. O que pode resultar de motivos tão simples quanto o facto de se receberem ao longo do dia mais notificações (se estas estiverem ativadas), incentivando a que se consulte mas vezes o grupo em causa. Em contrapartida, pode também dar-se o caso de não haver correspondência nos dias entre cada um destes quartis, o que invalidaria de imediato a hipótese. Daí que mais abaixo a hipótese seja reformulada para contemplar uma leitura mais aprofundada, a partir da métrica de envolvimento (*engagement*).

Tabela 1. Resumo estatístico das atividades (*posts*) e reações (*likes* e comentários) dos membros. [TCTVT]

<i>Quantidade de Posts = 982</i>	<i>Posts</i>	<i>Comentários</i>	<i>Likes</i>
Média	25,79	63	307
p25 (Quartil inferior)	21	25	207
p50 (Mediana)	26	48	311
p75 (Quartil superior)	33	68	395
p90 (Decil superior)	36	94	520
Valor máximo	43	606	602

Se tomados o decil superior (em particular no caso dos *likes*) e os valores máximos (em particular no caso dos comentários) para o período em análise, salta à vista uma anomalia, ou melhor, um possível *outlier*. E de facto, o dia com o maior número de publicações é também aquele em que se verificou o maior número de comentários e o segundo maior número de *likes*: 17 de junho de 2015. Nesse dia, o tema havia sido «filmes que vimos no cinema e do qual [sic] saímos a meio [...] ou] que, se eu soubesse o que sei hoje, tinha saído a meio»¹⁶, claramente apelando às diferentes experiências e sensibilidades estéticas dos membros do grupo. Tal deixa em aberto o que nesta fase nos limitaremos a enunciar como *Hipótese B*: assumindo uma categorização binária entre temas «factuais» – isto é, onde é possível aplicar algum tipo de teste para conferir se a publicação cumpre o estipulado¹⁷ – e «subjectivos», estes últimos incitam a maior envolvimento. Deixe-se contudo como salvaguarda (*Hipótese B1*) que este acréscimo no grau de envolvimento pode dar-se muito mais ao nível dos comentários do que dos *likes*: por exemplo, no período analisado verifica-se que o dia com maior número de *likes* é a véspera, 16 de junho, data em que o tema era claramente «factual», «filmes baseados em acontecimentos verídicos», e o número de comentários foi 217.

16. Cf. www.facebook.com/groups/tcdtevt/permalink/914365291935006/ (acesso restrito a membros do grupo). Daqui em diante, por questões de simplicidade, sintetizamos a descrição do tema desde que isso não leve a ambiguidade na sua interpretação.

17. Por exemplo por confronto com o título ou com a letra da música publicada, com uma cena ou sequência de *videoclip* (no caso de músicas) ou de filme, com a lista de nomes de atores ou personagens dum filme, etc.

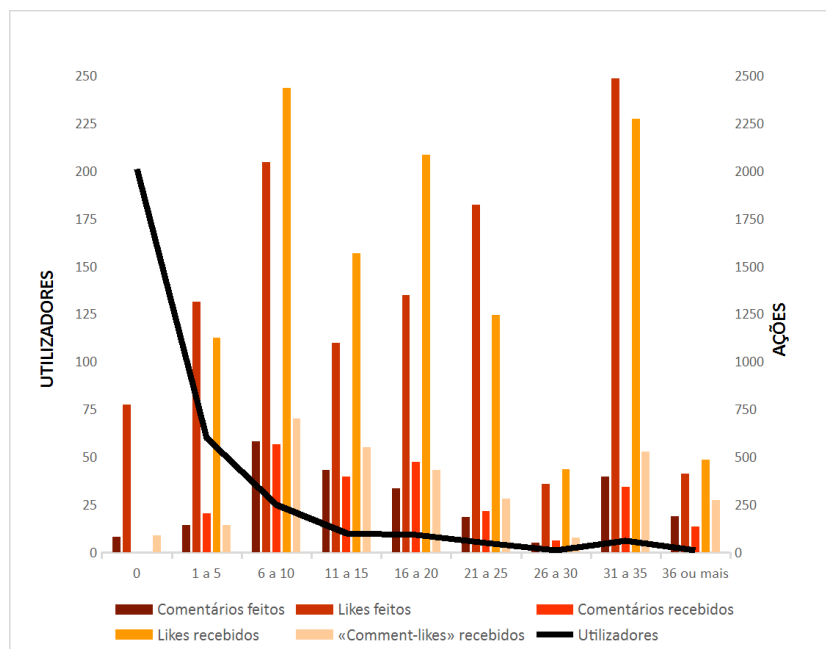


Figura 1. Resumo estatístico da atividade dos membros. [TCTVT]

A Figura 1 segue uma abordagem um pouco diferente. A curva a negro mostra a quantidade de utilizadores¹⁸ que fizeram *posts* durante o período analisado; temos, por exemplo, 201 que não fizeram nenhuma publicação (e que poderíamos classificar como «participantes passivos», passe o oxímoro), 60 que fizeram entre 1 e 5 publicações, etc. Para este primeiro indicador, mau grado ligeiras flutuações, temos o esperado: quanto maior este tipo de participação, menos os utilizadores, e esse decréscimo assume uma variação que praticamente coincide com a conhecida curva da «lei de potência» (cf. Barabási, 2003)¹⁹.

As barras mostram de seguida, para os mesmos agrupamentos de utilizadores, que outras atividades lhes estão associadas – não só o número de *likes* e de comentários feitos, mas também os recebidos (incluindo os *likes* a comentários). Os valores encontrados são de difícil interpretação devido à irregularidade do gráfico, sendo ainda mais difícil dar-lhes sentido através de uma hipótese para além do óbvio. E o óbvio é que «gostar» tende a ser a forma de interação dominante, e ainda que podem identificar-se pelo menos os seguintes tipos de comportamento: há membros que publicam pouco ou nada embora exerçam outras atividades (e por consequência também recebem *likes* ou menos frequentemente comentários), há membros que têm uma ativi-

18. Ignorando os mais de 2 mil membros que estiveram totalmente inativos.

19. E isto apesar da condicionante de apenas poder ser feita uma publicação por dia (à exceção dos administradores que podem fazer duas publicações se uma delas for a que lança o tema diário). Cf. as regras pelas quais se rege o grupo, disponíveis em www.facebook.com/notes/tribo-cin%C3%A9fila-das-tapas-vinhoto/regras/202826919755517 (acesso restrito aos membros do grupo).

dade bastante intensa a todos os níveis²⁰, e como é óbvio os que publicam com uma frequência moderada (as três colunas entre «11 a 15» e «21 a 25»). Menos óbvio, mas talvez mero fruto de uma singularidade na amostra recolhida, é o facto de nestas colunas intermédias haver um decréscimo de atividade (em especial *likes* feitos e recebidos) por comparação com a coluna «6 a 10», voltando a haver decréscimos – e ainda mais acentuados – mais à frente.

Veremos de seguida, ao observar o outro grupo, se aí há lugar para alguma outra hipótese de trabalho.

3.2. A Igreja Universal dos Fazedores de Bonitas Listas Musicais dos Últimos Dias

Se a atividade fosse proporcional ao número de membros, a média de publicações diárias da «igreja» deveria rondar as 145, os comentários as três centenas e meia, e os *likes* seriam mais de 1700. A realidade é contudo um pouco mais modesta. Em vez de 145, contabilizamos uma média de 100 publicações (quase o mesmo para o valor mediano); em vez de 350 comentários apenas 200 (166, no caso da mediana); e 1590 *likes* (1366 na mediana). Significa isto que se publica comparativamente menos do que na «tribo», e se comenta ainda menos, mas a quantidade de *likes* não decresce tanto, o que parece apontar – mas não façamos já disto hipótese – para um envolvimento mais «frouxo» neste grupo.

Sigamos esta pista, mas sem descurar os outros indicadores. Olhando para os quartis destas formas de atividade (Tabela 2), a sua amplitude é também maior: não tanto nos *posts* (mas ainda assim há mais variação do que no outro grupo), nos comentários é já um pouco mais significativa, mas é sobretudo nos *likes* que essa variação se faz sentir, contabilizando-se 858 no quartil inferior, e 2091 no superior (bem acima do dobro do primeiro quartil, quando na «tribo» se estava abaixo). O decil superior, com 2733 *likes*, e principalmente o *outlier* que é o valor máximo com 4114 confirmam esta leitura para a amostra analisada. A *Hipótese A* (cf. acima) não se vê afetada, sendo inclusive reforçada, embora neste grupo pareça haver um predomínio dessa forma mais descomprometida de envolvimento que é «gostar» duma publicação.

Tabela 2. Resumo estatístico das atividades (*posts*) e reações (*likes* e comentários) dos membros. [IUFBLMUD]

<i>Quantidade de Posts = 3 832</i>	<i>Posts</i>	<i>Comentários</i>	<i>Likes</i>
Média	100,42	203	1 590
p25 (Quartil inferior)	71	116	858
p50 (Mediana)	102	166	1 366
p75 (Quartil superior)	124	294	2 091
p90 (Decil superior)	162	380	2 733
Valor máximo	229	545	4 114

Uma análise mais aprofundada, que não cabe nestas linhas, poderá também esclarecer quanto à *Hipótese B* (se os temas mais factuais são os que geram mais envolvimento), embora neste

20. Não é de pôr de parte que seja nesta categoria que se concentram os administradores, *insight* cuja confirmação nos é vedada porque a API os «anonimiza».

grupo pareça ser mais lógica uma sub-hipótese alternativa (*B2*) segundo a qual o acréscimo no grau de envolvimento se dá muito mais ao nível dos *likes* do que dos comentários. A verificar-se que *B1* prevalece no grupo sobre cinema e *B2* no grupo sobre música, essa poderá ser uma diferença a reter.

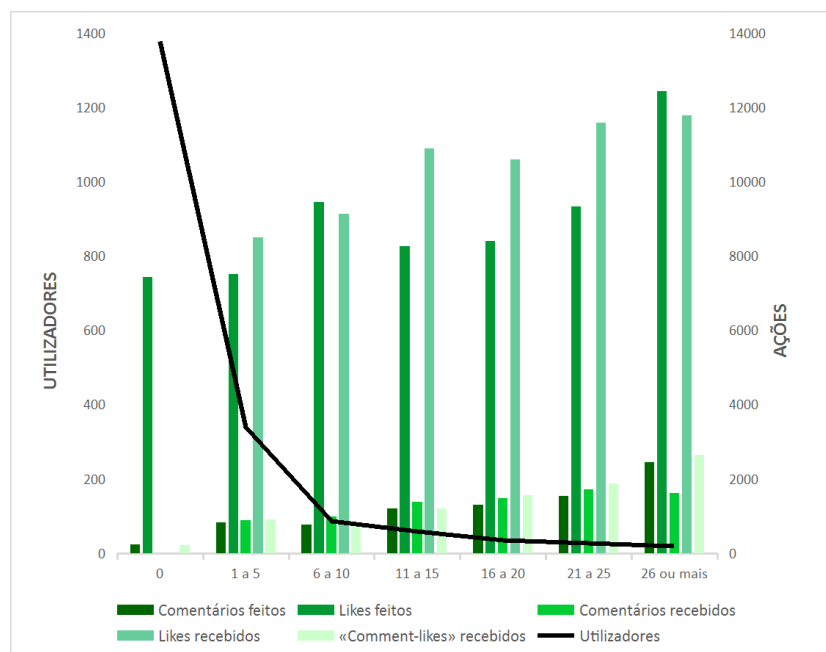


Figura 2. Resumo estatístico da atividade dos membros. [IUFBLMUD]

No que diz respeito à distribuição da atividade dos membros por diferentes classes (e já excluídos aqueles que não participaram de todo), reencontramos a habitual «lei de potência» e a sua curva a descrever a variação na participação destes relativamente à quantidade de publicações. Dos 1944 membros que tiveram algum tipo de atividade, 1378 apenas fizeram *likes* ou comentários, 339 publicaram uma a cinco vezes, e assim por diante, com os valores em cada classe sempre a diminuir até aos 19 que publicaram 26 ou mais vezes. Contudo, e ao contrário do grupo anterior, verifica-se uma regularidade muito maior nas outras atividades possíveis: ainda que com flutuações, a tendência é para que os membros que mais frequentemente publicam sejam também aqueles que mais *likes* e comentários fazem, e igualmente aqueles que mais *likes* e comentários recebem. Precisamos portanto, para aprofundar desta primeira visão geral, de dirigir o nosso olhar para a forma como estas atividades se traduzem na métrica conhecida como *engagement* ou envolvimento.

4. Análise do envolvimento («engagement») gerado

Começamos também aqui por abordar a atividade da TCTVT, seguindo-se a da IUFBMUD. E podemos fazê-lo com uma constatação algo trivial, mas que permite comprovar a aceitação dos membros das normas, no essencial similares em ambos os grupos, pelas quais se regem (cf. Rosa, 2015). Em cada dia propõe-se um tema, e a participação deve traduzir-se na publicação de um *link* para um vídeo com o *trailer* ou alguma cena específica dum filme (no primeiro caso) ou para o vídeo de uma música (no segundo) que satisfaça esse constrangimento temático.

Não é portanto de estranhar que, em 982 publicações, 925 (94,2%) sejam ligações para o YouTube ou para domínios alternativos deste site (youtu.be e m.youtube.com²¹), com ocasionais ocorrências para domínios como o imdb.com, onde também se encontram *trailers* e excertos de filmes. Todas as outras possibilidades – em particular fotos e *links* para outros domínios – são residuais, ressaltando-se o caso das atualizações de estado, 42 no total, que na sua maioria correspondem às publicações que lançam o tema diário (ainda que por vezes também se usem ligações a fotos para esse efeito). O que resta são quase certamente *posts* que não cumprem as regras mas que por alguma razão não foram apagados pelos administradores²².

Estabelecida esta constância – quase diríamos monotonia – na atividade do grupo, e portanto posta de parte (excepto no que respeita ao *post* de lançamento do tema, como veremos já de seguida) qualquer correlação entre tipos de publicação e envolvimento, voltamos a nossa atenção para uma outra variável que poderá afetar esse envolvimento: os utilizadores.

Definimos como envolvimento («engagement») o somatório quantitativo das ações que podem ser tomadas para cada publicação: *likes*, comentários, e *likes* a comentários²³. Uma análise aprofundada ao grau e tipo de envolvimento em cada um dos grupos exigiria uma descrição dia a dia, que aqui não poderá ser encetada por questões de espaço²⁴. Ainda assim, será útil observá-lo para a totalidade do período.

Tabela 3. *Posts* e respetivo envolvimento (decomposto também em «gostos», comentários, e «gostos» aos comentários). [TCTVT]

	<i>Posts</i>	<i>Likes</i> ²⁵	Comentários	<i>Likes</i> a comentários	Envolvimento
Total	982	11.654 (11.688)	2.383 (2.392)	3.067	17.095
Imagem	5	54 (54)	24 (24)	23	101
<i>Link</i>	39	476 (479)	77 (77)	87	640
<i>Status</i>	42	337 (337)	196 (170)	253	759
Vídeo	896	10.778 (10.819)	2.113 (2.121)	2.704	15.595

21. Que na tabela 3 não são reconhecidos como vídeos, e daí a aparente incongruência nos valores.

22. Cf. de novo Rosa, 2015 para uma explicação deste procedimento, embora restrita ao grupo de que falaremos adiante.

23. Só muito recentemente, num momento posterior ao período aqui em análise, foi introduzida nos grupos a funcionalidade dos «threaded comments», isto é, a de comentar em comentários, criando uma sub-ramificação.

24. À exceção de um dia atípico mas, como veremos, revelador.

25. Nesta coluna e na seguinte, surgem em cada célula dois valores. O primeiro regista a contagem no momento do rastreamento, correspondendo ao número remanescente, ora de *likes* ora de comentários. O segundo, entre parêntesis,

Para a «tribo» verifica-se, como seria de prever, que o maior contributo para o envolvimento – aproximadamente dois terços do total – vem dos *likes*, a forma mais simples e superficial de interação, quer tomemos o conjunto das publicações (cerca de 11,9 *likes* por publicação) ou aquelas que mais definem a razão de ser do grupo, os vídeos (12 *likes* por publicação). Mas é nos comentários que vale a pena concentrarmo-nos. Em média (cf. Tabela 3 para os valores absolutos), há cerca de 2,4 comentários por cada post (um pouco menos, 2,36, se tomados apenas os vídeos), o que, a não ser que seja contrariado por uma análise mais «fina», dia a dia ou mesmo *post a post*, leva a crer que – *Hipótese C* – a atividade do grupo não é geradora de diálogo consistente e continuado. No entanto, deve ser destacado o pormenor de esta média de comentários por *post* subir para cerca de 4,6 no caso das publicações de tipo *status*. Mesmo sem olhar à lupa para estes casos, podemos afirmar com bastante segurança, devido à participação frequente no grupo, que estas são quase exclusivamente as que lançam o tema diário, o que leva a crer – *Hipótese C1* – que o diálogo seja nessas situações mais significativo, entre elogios ou críticas à escolha do tema, pedidos de esclarecimento caso este apresente ambiguidades, etc. Olhemos para a «igreja» para confirmar se estas hipóteses são também aí aplicáveis.

Nesse outro grupo, a IUFBLMUD, encontramos também um notório predomínio dos *posts* de tipo vídeo: 3750 em 3872 (96,8%) do YouTube e seus domínios alternativos, e isto sem contar com algumas ligações ocasionais que a API do Facebook não contabiliza como vídeos embora apontem para este tipo de conteúdo (dailymotion.com, mojvideo.com, e outros). Notório e previsível, pois, de acordo com as regras do grupo, tudo o que resta é anómalo a não ser que se trate da publicação que apresenta o tema diário – como no grupo anterior, quase sempre *status*, ainda que por vezes possa ser uma foto ou uma ligação. Também neste grupo o maior contributo para o envolvimento é proveniente dos *likes*, seja qual for o tipo de publicação. A média de *likes* por *post* é de cerca de 15,8, quer no geral quer restringindo aos vídeos, o que revela um comportamento bastante coerente com o do outro grupo²⁶.

Tabela 4. *Posts* e respetivo envolvimento (decomposto também em «gostos», comentários, e «gostos» aos comentários). [IUFBLMUD]

	<i>Posts</i>	<i>Likes</i>	Comentários	<i>Likes</i> a comentários	Envolvimento
Total	3832	60.399 (60.532)	7.731 (7.736)	8,645	76,775
Evento	1	0 (0)	0 (0)	0	0
Imagem	9	165 (165)	57 (57)	53	275
<i>Link</i>	63	857 (863)	130 (130)	153	1,140
<i>Status</i>	49	668 (668)	263 (263)	311	1,242
Vídeo	3,710	58.709 (58,836)	7.281 (7.286)	8,128	74,118

Se tomado apenas o contributo dos comentários, a média geral é de 2,29 por publicação, baixando para 1,96 se tomadas apenas as de tipo vídeo, o que é consistente com a Hipótese C formulada acima. Ora, apesar do carácter pouco frequente das publicações de tipo *status*, é aí, corresponde à quantidade total registada. Optámos neste caso pelo primeiro valor para os nossos cálculos de médias, percentagens e rácios.

26. Sendo a média ligeiramente superior explicável (ainda que não seja de todo proporcional) pelo maior número de membros e pelo maior grau de atividade.

tal como no grupo anterior, que se encontra uma maior densidade de comentários: cerca de 5,34 por *post* para ser mais preciso, igualmente deixando de pé a Hipótese C1. Esta média sobe ainda mais no caso das fotos (6,3), mas sendo estes casos raros (e quase sempre variações na forma de apresentar o tema diário), não é pertinente alterar a hipótese.

Vale ainda a pena, embora com uma função mais tangencial a esta nossa primeira aproximação ao funcionamento destes grupos, detetar os dias em que se verificou maior grau de envolvimento e também – havendo ou não coincidência no dia – os *posts* que se destacaram como *outliers*²⁷. Começamos desta vez pela «igreja», cujos três picos de maior envolvimento se dão, respetivamente, nos dias 6 de julho (com o tema «palavra “não” no título da música»²⁸, 4666 de *engagement*), 23 de junho («ao meu lado, no título ou na letra», com uma contagem de 4338), e 10 de junho («nomes de bebidas alcoólicas, no título ou na letra», com 4208). Todos temas «factuais» (cf. a sua definição acima) e de certa forma «fáceis»²⁹, embora não consideremos estar perante dados suficientes para fazer disto hipótese. Já no caso da «tribo», os picos têm lugar a 17 de junho («filmes vistos no cinema e do quais saíram ou deveriam ter saído a meio», com 1616 de *engagement*), a 16 de junho («filmes baseados em acontecimentos verídicos», com 805), e 30 de maio («filme com que se identificam, e que já viram mais de mil vezes», com 760). Aqui, e ao contrário do grupo dedicado à música, apenas um destes picos ocorre com um tema factual³⁰, o que parece apontar – mas, de novo, não o suficiente para gerar uma hipótese – para um maior incentivo à participação nestes casos que potencialmente podem gerar maior bipolaridade nas opiniões. Note-se ainda que este maior envolvimento não replica a variação de publicações diárias; isto é, não foram dias de pico de publicações (ainda que acima da média), e sim nas reações a estas³¹. Também aqui o comportamento dos membros da «tribo» diverge do da «igreja», pois nesta parece haver uma tendência mais firme para que os dias de maior *engagement* sejam também aqueles em que há maior número de posts (com coincidências, embora numa ordem diferente, no «top 5» de publicações e de envolvimento³²). Ainda que a amostra possa ter ficado enviesada por algum tipo de atipicidade, esta divergência parece-nos suficiente

27. Sendo certo que seria igualmente relevante tomar a abordagem inversa no caso dos temas – quais os que geraram menos *engagement* –, é tarefa que terá de ficar para ocasião futura. Relativamente a *posts* individuais, trata-se de algo que só faria sentido se fosse possível identificar os utilizadores que as publicaram (se têm muitas ou poucas conexões de amizade dentro do grupo, etc.), algo que a API do Facebook não permite.

28. Cf. a nota 17.

29. Por razões distintas. No primeiro caso porque, embora limitando ao título, é uma palavra frequente; no segundo porque, em particular no universo musical anglo-saxónico, é um lugar-comum em canções de amor; no terceiro pelo amplo leque de escolhas.

30. Sendo certo que seria igualmente relevante tomar a abordagem inversa no caso dos temas – quais os que geraram menos *engagement* –, é tarefa que terá de ficar para ocasião futura. Relativamente a *posts* individuais, trata-se de algo que só faria sentido se fosse possível identificar os utilizadores que as publicaram (se têm muitas ou poucas conexões de amizade dentro do grupo, etc.), algo que a API do Facebook não permite.

31. 17 de junho em 4.º lugar no número de *posts*, *ex-aequo* com 3 de julho; 16 de junho em 3.º; 30 de maio em 9.º, *ex-aequo* com 11 de junho e 6 de julho. O «top 5» em número de *posts* corresponde, por ordem decrescente, aos dias 1 de junho (7.º em envolvimento), 4 de junho (4.º em envolvimento), 16 de junho (2.º em envolvimento), e, *ex-aequo*, 17 de junho (1.º em envolvimento) e 3 de julho (21.º em envolvimento).

32. 6 de julho (1.º em *posts*, 2.º no envolvimento), 9 de junho (2.º em *posts*, 5.º no envolvimento), 3 de junho (3.º em *posts*, 4.º no envolvimento), 23 de junho (4.º em *posts*, 1.º no envolvimento), 10 de junho (5.º em *posts*, 3.º no envolvimento).

Dito de outra forma, como temos vindo a verificar, se a esmagadora maioria dos membros do grupo não participa de todo, sendo a dinâmica destes grupos assegurada por uma minoria, o mesmo parece também repetir-se, numa outra escala, entre utilizadores que, ainda que ativos, se envolvem pouco, e a ainda mais pequena minoria que – mesmo que de forma seletiva (em determinados dias, em determinados *posts*, ...) – assegura a continuidade da coesão do grupo. Um outro tipo de análise e visualização da rede, terminando este percurso, poderá ajudar-nos a reforçar essa intuição.

5. Desigualdades na rede

Uma abordagem complementar à que temos vindo a seguir, permitindo conciliar uma perspetiva global com a identificação de comportamentos mais individualizados³⁴ resulta da possibilidade de materializar a rede de interações em grafos, de acordo com diferentes parâmetros – métricas como o *indegree*, o *outdegree*, o envolvimento, etc. – e algoritmos de visualização. Para o nosso objetivo, que consiste em perceber como agem os membros destes dois grupos relativamente às ações permitidas pela plataforma do Facebook, e depois de experimentadas diversas combinações de parâmetros e algoritmos, nomeadamente a que deu origem à Figura 3, considerámos que a que melhor ia ao encontro do objetivo e das hipóteses que foram emergindo das análises acima era a que passamos a descrever.

Como vimos já, a maior contribuição para o envolvimento provém invariavelmente do ato de «gostar», a forma mais superficial de interagir. Na medida em que isso pode enviesar a leitura, ignorámos o contributo dos *likes*, concentrando-nos então nas outras duas – *posts* e comentários –, procurando assim perceber como se posicionavam os membros destes grupos relativamente a cada uma destas ações.

O algoritmo usado foi o «Circular Layout»³⁵, tomando os utilizadores como nodos, que foram parametrizados de acordo com essas duas variáveis (número de publicações e número de comentários). As linhas – direcionadas, como descrito acima – representam as formas de envolvimento («gostar» ou comentar), sendo estas mais espessas quanto maior for esse grau de envolvimento entre dois utilizadores concretos. O elemento mais importante para uma adequada interpretação dos grafos tem contudo a ver com o tamanho e com a posição dos nodos (isto é, o modo como estão ordenados ao longo da circunferência que representa a rede). Quanto ao tamanho, estes são tanto maiores (e simultaneamente mais escuros) quanto maior o número de publicações do utilizador em causa³⁶; quanto à sua ordenação ao longo da circunferência externa, esta depende do outro contributo para o *engagement* que considerámos relevante: a quantidade de comentários feitos.

Isto é, imaginando uma situação ideal em que os utilizadores que mais comentam fossem também os que mais publicam, e pela mesmíssima ordem, o grafo daí resultante seria algo como

34. Apesar do anonimato dos utilizadores que é imposto pela API do Facebook.

35. Utilizado para perceber, de forma ordenada, a distribuição dos nodos e respectivas ligações; no Gephi tem de ser carregado como *plugin*.

36. Aqueles que não publicaram mas fizeram *likes* ou comentários são apenas pontos (com uma ligeira espessura para melhor identificação) ao longo da circunferência.

o que mostra a figura 4³⁷: uma total redundância entre espessura, tom de cor e posição na circunferência.

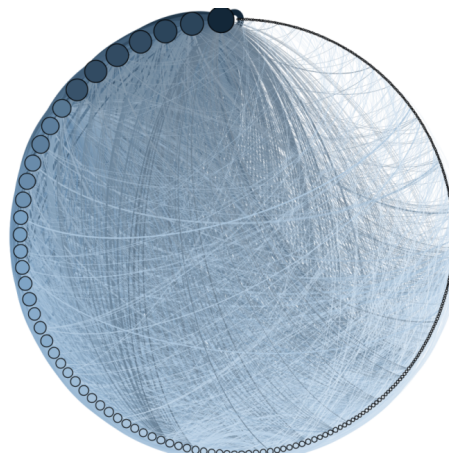


Figura 4. Rede ideal de publicações, comentários e envolvimento.

Em vez disso, os grafos – respetivamente para a «tribo» e para a «igreja» – tomaram, para o período analisado, o aspeto que pode ver-se nas Figuras 5 e 6, e que nos leva a formular a nossa última hipótese de trabalho. *Hipótese D*: Entre os utilizadores ativos de ambos os grupos (e desprezando o contributo dos *likes*), há uma tendência para que estes se distribuam entre «publicadores», «comentadores», e aqueles que, pertencendo a ambos os subconjuntos, podem ser classificados como «publicadores-comentadores». Ainda que as diferenças entre ambas as figuras possam não ser significativas, será ainda cauteloso propor uma *Hipótese D1*: essa interseção de «publicadores-comentadores» aparenta ser maior no caso da TCTVT, que pelo menos segundo estas métricas (número de *posts* e número de comentários por utilizador) poderá ser considerado um grupo mais coeso.

37. Criado usando a rede da «tribo» e modificando os parâmetros de forma a ordenar por número de publicações e não de comentários.

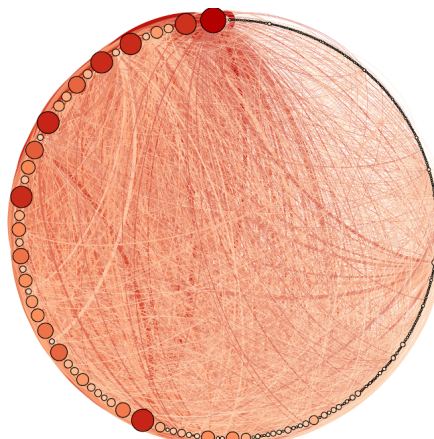


Figura 5. Rede de publicações, comentários e envolvimento. [TCTVT]

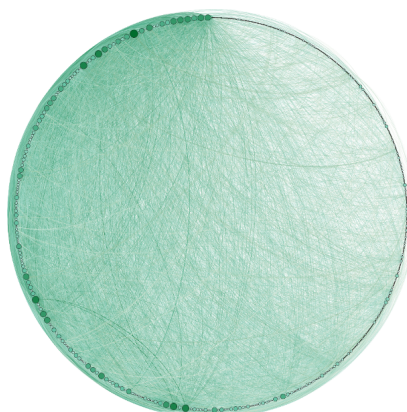


Figura 6. Rede de publicações, comentários e envolvimento. [IUFBLMUD]

6. Síntese conclusiva

Num estudo que se assume como uma análise exploratória de dados, seria prematuro falar de conclusões. É certo que não se partiu de um desconhecimento absoluto, dada a familiaridade prévia relativamente ao quotidiano de cada um destes grupos, mas, como o dita a necessidade científica de romper com o senso comum (mesmo que para mais tarde confirmá-lo, em parte ou no todo), a informação recolhida e o que ela permite revelar num sentido quantitativo confrontam-nos com a necessidade de proceder à criação de um conjunto de hipóteses que possam, em futuras pesquisas, guiar-nos de forma mais segura.

Encerramos por isso este artigo com uma síntese das hipóteses que foram sendo sugeridas pelas análises acima apresentadas.

Hipótese A: Tende a haver um efeito de *feedback* positivo entre o número de publicações num determinado dia e o envolvimento para esse mesmo período: quanto mais *posts*, maior o envolvimento para cada um dos *posts* nesse mesmo dia.

Hipótese A1: Especificamente para estes dois grupos, ainda que um maior número de *posts* num dia tenda em ambos os casos a gerar maior envolvimento para cada um dos *posts*, e portanto se some maior envolvimento nesse mesmo dia, a tendência é mais significativa no caso do grupo IUFBLMUD do que na TCTVT.

Hipótese B: Assumindo uma categorização binária entre temas «factuais» e «subjectivos», estes últimos geram, em ambos os grupos, maior envolvimento.

Hipótese B1: Este acréscimo no grau de envolvimento dá-se, no caso da TCTVT, muito mais ao nível dos comentários do que dos *likes*.

Hipótese B2: Este acréscimo no grau de envolvimento dá-se, no caso da IUFBLMUD, muito mais ao nível dos *likes* do que dos comentários.

Hipótese C: A atividade destes grupos não é geradora de diálogo consistente e continuado.

Hipótese C1: Ainda assim, diálogo é mais significativo nas publicações que lançam o tema do que nas que correspondem à atividade mais comum (publicar *links* para vídeos de acordo com o tema diário proposto).

Hipótese D: Entre os utilizadores ativos de ambos os grupos (e desprezando o contributo dos *likes*), há uma tendência para que estes se distribuam entre «publicadores», «comentadores», e aqueles que, pertencendo a ambos os subconjuntos, podem ser classificados como «publicadores-comentadores».

Hipótese D1: Essa interseção de «publicadores-comentadores» aparenta ser maior no caso da TCTVT, que pelo menos segundo estas métricas (número de *posts* e número de comentários por utilizador) poderá ser considerado um grupo mais coeso.

Trata-se, recordemo-lo, de hipóteses que se aplicam exclusivamente aos grupos que nos propusemos estudar, ainda que acreditemos que possam servir de primeiro guia a outros investigadores que pretendam, como foi o presente caso, concentrar-se na dinâmica dos grupos do Facebook, por contraste por exemplo com as páginas institucionais. Reservamos para esse caso uma última hipótese, de carácter mais global, e que formulamos assim.

Hipótese global: Parece haver nestes grupos, salvaguardada a excepção dos *outliers*³⁸, um tipo de interação diferente da que é associada às páginas³⁹. Ainda que em ambos os casos a facilidade – ou mesmo descomprometimento – dos *likes* não deixe de ser a modalidade dominante, quando nos focamos nos comentários é de supor que estes gerem uma dinâmica mais «intensa» (i. e., longas discussões entre múltiplos utilizadores) nas páginas e mais «dispersa» (i. e., mais condensados) nos grupos. Em contrapartida, haverá nos grupos uma tendência para maior regularidade na interação, que se liga a uma sensação de pertença, por contraste com uma interação mais pontual e distanciada no caso das páginas.

38. Como o que isolámos acima, no final do ponto 4.

39. Cf. o artigo «15 de Março: “O Brasil foi pra Rua! – De Novo!”», dos mesmos autores e neste mesmo volume.

Referências

- Adamic, L. (2014). *Social network analysis*. Disponível em: www.coursera.org/course/sna
- Barabási, A.-L. (2003). *Linked: how everything is connected to everything else and what it means for business, science, and everyday life*. Nova Iorque: Basic Books.
- Bastian M.; Heymann, S. & Jacomy, M. (2009). Gephi: an open source software for exploring and manipulating networks. *International AAAI Conference on Weblogs and Social Media. From AAAI* [PDF]. Disponível em: www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154
- Jacomy, M.; Heymann, S.; Venturini, T. e Bastian, M. (2012). ForceAtlas2: A Continuous Graph Layout Algorithm for Handy Network. *PLoS ONE*, 9(6), doi:10.1371/journal.pone.0098679
- Omena, J. J. C. & Rosa, J. M. (2015). Estudos no Facebook em Portugal: revisão sistemática dos métodos de investigação. *Communication Studies* (Número especial: Advanced Methodologies in Communication Studies), 18: 15-33. Disponível em: www.ec.ubi.pt/ec/19/pdf/n19a02.pdf
- Rieder, B.; Abdulla, R.; Poell, T.; Woltering, R. & Zack, L. (2015). Data critique and analytical opportunities for very large Facebook pages: lessons learned from exploring “We are all Khaled Said”. *Big Data & Society*, Julho-Dezembro: 1-22. doi: 10.1177/2053951715614980
- Rogers, R. (2013). *Digital methods*. Cambridge (MA): The MIT Press.
- Rosa, J. M. (2015). Devoções. *Novos Olhares: Revista de Estudos Sobre Práticas de Recepção a Produtos Midiáticos*, Setembro, 4(1): 165-177. doi: 10.11606/issn.2238-7714.no.2015.102232
- Tukey J. W. (1977). *Exploratory data analysis*. Reading (MA): Addison-Wesley.